

- LEE, S. W., S. C. KIM, W. W. KIM, S. D. HAN and K. B. YIM (1997): Characteristics of leaf morphological, vegetation and genetic variation in the endemic populations of a rare tree species, *Koelreuteria paniculata* Laxm. *Jour. Kor. For. Soc.* **86**: 167–176.
- LEE, S. W., S. C. KIM and H. S. LEE (1998): Allozyme variation in *Abeliophyllum distichum* Nakai, an endemic tree species of Korea. *Silvae Genetica* **47**: 294–298.
- LEE, S. W., W. W. KIM, B. C. LEE, Y. Y. KIM and S. C. KIM (1995): Genetic variation of acorn production stands in *Quercus acutissima* and *Q. variabilis*. *Kor. Jour. Breed.* **27**: 345–358.
- LESICA, P., R. F. LEARY, F. W. ALLENDORF and D. E. BILDERBACK (1995): Lack of genetic diversity within and among populations of an endangered plant, *Howellia aquatilis*. *Conservation Biology* **2**: 275–282.
- LOVELESS, M. D. and J. L. HAMRICK (1984): Ecological determinants of genetic structure in plant populations. *Ann. Rev. Ecol. Syst.* **15**: 65–95.
- NEL, M. (1978): Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* **89**: 583–590.
- SNEATH, D. L. and R. R. SOKAL (1973): *Numerical Taxonomy*. WH Freeman and Co., San Francisco, USA. 573pp.
- SOKAL, R. R. and F. J. ROHLF (1981): *Biometrics*. 2nd ed. WH Freeman and Co., San Francisco, USA. 859p.
- SWOFFORD, D. L. and R. B. SELANDER (1989): BIOSYS-1: a computer program for the analysis of allelic variation in population genetics and biochemical systematics. Release 1.7. Illinois Natural History Survey.
- WALLER, D. M., D. M. O'MALLEY and S. C. GAWLER (1987): Genetic variation in the extreme endemic *Pedicularis furbishiae* (Scrophulariaceae). *Conservation Biology* **1**: 335–340.
- WEEDEN, N. F. and J. F. WENDEL (1989): Genetics of plant isozymes. *In: Isozymes in Plant Biology*. Edited by D. E. SOLTIS and P. S. SOLTIS. Dioscorides Press. Portland, Oregon, USA. pp.46–72.
- WON, B. O. (1981): *Illustrated Flora and Fauna of Korea* Vol. 25 Avifauna. Ministry of Education.
- WRIGHT, S. (1978): *Evolution and the genetics of population. variability within and among natural population*. University of Chicago Press. Chicago, USA.

Comparison of Phenotype and Combined Index Selection at Optimal Breeding Population Size Considering Gain and Gene Diversity

By H. LI^{1,*} and D. LINDGREN²⁾

(Received 1st July 2005)

Abstract

A breeding program was simulated in this study. Two alternative ways of selecting the breeding population for the following generation was compared. Phenotypic selection, which means to select just on the individual performance, and combined index selection, which means selection on predicted breeding value for each individual obtained by weighting family average and individual phenotype, were compared. The plant number (testing resource) and gene diversity (status number, N_s) were kept constant, but the breeding population size was variable and chosen for maximizing gain for the particular breeding scenario. At low and medium heritability phenotypic selection was inferior to combined index selection. Only when heritability was high phenotypic selection was as efficient (generation 1) as or more efficient (generation 5) than combined index selection. This contrasts to earlier studies done under constant breeding population size, where selection methods appeared similar. The advantage in gain of combined index selection is usually at a larger breeding popula-

tion size. At limited heritability and breeding population size the difference is considerable. When breeding population size was kept rather small (<100), and the heritability limited, combined index selection can result in slightly higher gain than phenotypic selection at the same gene diversity, but this was at the cost of a much larger breeding population. Phenotypic selection and combined index selection appears as rather similar for many cases in this simple model used in this study. Considering other advantages with phenotypic selection, it may often be regarded as a competitive alternative.

Key words: heritability, genetic gain, status number, gene diversity, breeding population size, selection efficiency.

Introduction

Genetic gain and gene diversity are two goals which plants breeders want to combine, it means that a compromise between these two goals has to be made. Phenotypic selection (PS, mass selection) is based on individual performance only. Combined index selection (CIS) means selection based on an index combining individual values with family averages weighting the two sources of information to maximise the correlations between the index and the true breeding values of the individuals, and it maximizes genetic gain (FALCONER, 1989). Phenotypic selection (PS) is a classical method, it is very sim-

¹⁾ College of Forest Resources and Environments, Nanjing Forestry University, Nanjing 210037, P. R. China.

²⁾ Department of Forest Genetics and Plant Physiology, Swedish University of Agricultural Sciences, S-901 83 Umeå, Sweden. Email: Dag.Lindgren@genfys.slu.se.

^{*} Corresponding Author: HUOGEN LI, Tel: 0086-25-85428731, Fax: 0086-25-85427412, email: hgli@njfu.edu.cn.

ple and it can be considered cost efficient (BURTON, 1974 a, b; COTTERILL, 1986; DEMPFLE, 1990). Several papers have been published regarding genetic gain and gene diversity, WEI and LINDGREN (1991) reported that combined index selection (CIS) was inferior to phenotypic selection (PS) if compared at the same effective population size but in its comparison the selection intensity was allowed to vary and the family number and size were considered infinite. ANDERSSON et al. (1998) and others found that at high heritability phenotypic selection was as efficient as restricted combined index selection with restrictions on the number of offspring per parent selected, and not much inferior at low heritability. In such earlier comparisons family size and parent number (breeding population size) were constant. To control the level of gene diversity at combined index selection, restriction was made (e.g. constraining the maximum number of offspring a parent can have). Restrictions decrease the efficiency of selection and are generally unfair to comparisons where combined index selection is impaired with by introducing restrictions. The faster loss of gene diversity often occurring by combined index selection could be compensated by increasing breeding population size, but this has seldom or never been considered. The size of the breeding population can be optimized and comparing different breeding strategies at the same breeding population size may result in biased conclusions (DANUSEVICIUS and LINDGREN, 2005). To compare the efficiency of different selection methods, the gene diversity in term of status number (Ns), the overall cost should be kept constant while the breeding population size should be allowed to vary.

The aim of this study is to compare the efficiency of phenotypic selection versus combined index selection without restrictions on breeding population size at the same status number and at the same number of experimental plants.

Methods

The model breeding program

A tree improvement program can be thought of as repeated cycles with mating, testing and selection. The particulars of the simulated breeding program was chosen to facilitate a relevant comparison of phenotypic selection and index selection over a range of conditions, to be simple, to allow fast simulation, and to be relevant for real breeding. Selection was done without any restrictions either on the phenotype or on a combined index optimally weighting family and individual value. Mating design was double pair mating (DPM), each individual in the breeding population is the parent of two full sib families (selfing excluded), and parent pairs were formed randomly. DPM is a standard for Swedish conifer breeding (ROSVALL, 1999). Initially, individuals were assumed unrelated and not inbred, during the simulated breeding program various degrees of relationship and inbreeding developed. The family number was set to 5, 10, 20, 40, 80, 160, 320 and 640. To keep the total budget constant, the total number of test plants was fixed to 1280, so the corresponding family sizes were 256, 128, 64, 32, 16, 8, 4, and 2, respectively. The breed-

ing population size is the same as the number of selections, thus each selection is used as a parent for two families in the subsequent generation. The number of selections was the same as the number of families in order to keep the breeding population constant over time. The “budget” was independent on the number of selections. All genetic variance was assumed to be additive; dominance and epistatic variances were assumed to be absent. Initial genetic variance was preset, so the initial true narrow sense heritability was 0.01, 0.2, 0.5, and 1.0, but heritability changes over generations.

Simulation

The complexity, the high costs of each experimental tree and its handling, the complex interaction among individuals in field trials and the long time needed of multi-generation breeding program inevitably lead into simulation as a method of evaluation of different alternative breeding strategies. A software, the stochastic quantitative tree breeding simulator POPSIM, was used for this study (we used a version from 2002, an earlier version, which essentially is identical to the current, was described by MULLIN and PARK, 1995).

To improve accuracy of comparisons, the compared alternatives were initiated with the same founders when possible by using the same seed number to the simulator pseudorandom-generator.

The simulation accepts the size of the breeding population as entry. Individual simulations results in uneven values on status number which are different for the different selection strategies. To compare the two methods at exactly the same status number, linear interpolation (which works if the breeding population sizes chosen are close enough) between adjacent values was applied rather than using a time-consuming and cumbersome iterative process.

The status number was defined by LINDGREN et al. (1996) as half the inverse of the group coancestry (Θ , COCKERHAM, 1967):

$$N_s = \frac{1}{2\Theta} \quad (1)$$

Where group coancestry (Θ , which can be interpreted as the average coancestry in the population or a measure of gene diversity lost compared to “the wild forest”) is the probability that two genes, which are taken at random from the gene pool with replacement, are identical by descent:

$$\Theta = \frac{1}{n^2} \sum_i \sum_j \theta_{i,j} \quad (2)$$

where n is the population size, θ_{ij} is the coancestry between individual i and j .

Gain and status number (Ns) were calculated for a selection among progenies originated from double pair mating (DPM) crossings of a defined population size. In order to compare the difference in gain of the two methods among different gene diversity, individual effects were stochastically generated from normally distributed genetic and environmental effects with expected means and variances preset or derived by POPSIM. The total

phenotypic variance is the sum of additive genetic and environmental components. Note that the heritability was calculated by POPSIM as by a real breeder, who does not know the true value. Therefore there were small fluctuations in calculated heritability among iterations. If a real initial heritability was set to 1.0, POPSIM will act as if it were slightly below 1.0, as observations in iterations above 1 will be interpreted as 1.0, as a sensible breeder would do. Heritability sinks slightly from generation to generation as the genetic variance decreases over generations just the initial “true” heritability is shown. The number of iterations was chosen to 100, which turned out sufficient to get reproducible results. Simulation was carried out for five generations.

The superiority by phenotypic selection for genetic gain (δG) and breeding population size (δN) at the same status number can be formulated.

$$\delta G = (G_{PS} - G_{CIS})/G_{PS}; \quad \delta N = (N_{PS} - N_{CIS})/N_{PS}, \quad (3)$$

where G_{PS} and G_{CIS} are the gain of phenotypic selection (PS) and combined index selection (CIS) at the same status number respectively; N_{PS} and N_{CIS} are the breeding population sizes of phenotypic selection (PS) and combined index selection (CIS) at the same status number respectively.

Results

Gene diversity (status number) and genetic gain

Genetic gain decreased with increasing status number. The steepness of the decrease varied with selection strategy, heritability and generation (*Table 1*).

The selection methods and heritability

When initial heritability was very low (0.01), combined index selection was much more efficient than phenotypic selection, the superiority of combined index selection was 11–126% (generation 1) and 9–60% (generation 5) (*Figure 1*). At medium heritability (0.2, 0.5), still combined index selection appears more efficient than phenotypic selection, but as heritability increases, the superiority of combined index selection decreases and phenotypic selection seemed slightly better at generation 5 and $h^2 = 0.5$ (*Figure 1*). When heritability is high (0.5–1.0), phenotypic selection is superior or approximately equal to combined index selection. After 5 generations, the genetic gain is 0.6–17% higher, but at the first generation, there is no evident difference between phenotypic selection and combined index selection (*Figure 1*).

Accumulated genetic gain over generations

Phenotypic selection appeared better compared to combined index selection in the 5th generation than in the first (*Table 1*, *Figure 1*). It may be formulated that phenotypic selection appears to be more sustainable and relatively speaking more to its advantages in long-term breeding.

Gain and breeding population size

The size of the breeding population needed to achieve a certain status number is considerable lower for low

status number than for high, in particular for low heritability. The differences developed differently over generations for different heritability, for high heritability they become larger, but for low heritability slightly smaller (*Figure 2*). At low breeding population size (<100), considerably larger breeding population sizes were optimal for combined index selection than for phenotypic selection (unless heritability was not very high). At moderate heritabilities (0.2; 0.5) the alternatives resulted in similar gain, but with a much higher breeding population size at moderate population sizes (<100) (*Table 1*).

Discussion

The simulated breeding program versus real breeding programs

The model breeding program considered here was designed heading for a tree breeding program with typical components and covering a range of situations, but of course real programs are usually different to a varying degree. Full pedigrees with known seed parents and pollen parents for the breeding population are assumed. Simpler programs may use only seed parent pedigree. It seems that phenotypic selection is relatively more favourable when the pedigree is less controlled. The high efficiency of phenotypic selection is an argument to use breeding systems without known pedigrees (e.g. LINDGREN, 2003). All members in the breeding population contribute the same number of offspring. That was till recently the case for most tree breeding programs at least for segments of the breeding program, like an elite or nucleus population. There are advantages in structuring the breeding population with positive assortative mating and with allocating more resources to better parents (e.g. LSTIBUREK, 2005). The mating design (DPM) is typical for Sweden conifer breeding, it is suitable for principal studies as there are many parents as families, and it has been used as a case study for at least 20 papers. Our simulated program does not usually work with constraint (within family selection or a highest number of selected per family) although many earlier programs do, but these constraints make real programs less efficient compared to our model. Our program uses a small total resource, which may be kept in mind when drawing conclusions and probably favors phenotypic selection. However our scenarios cover a range of plants per family and include numbers considerable larger than most practical programs, and is thus even from that point of view rather realistic.

Unrestricted combined index selection and restricted combined index selection.

Restricted combined index selection here means that first combined index values were calculated for each individual, when the individuals were ranked for these values and selections were made from the top. But if a candidate for selection share the same parent as a number of higher ranking selections (= restriction limit), it is not selected. Phenotypic selection and restricted combined index selection can be compared by setting the restriction limit to the value, which gives the same status number (interpolation techniques were used).

Two cases were simulated at initial heritability 0.20 (Table 2). The entries were chosen to make it roughly compatible with some of the results obtained by ANDERSSON (1998b) to demonstrate the differences between the different concepts of comparing at the same status number. The difference between CIS and CISr was obvious, at generation 1 and status number 89, PS was 16.31% inferior to CIS but 7.87% superior to CISr; at generation 5 and status number 22, PS was almost as efficient as

CISr but was 7.31% inferior to CIS. CIS was 20.79% superior to CISr in the first generation and 6.11% in generation 5 (Table 2). Contrasting, ANDERSSON reported that phenotypic selection (PS) was almost as efficient at generation 1, or slightly less efficient at generation 5 as restricted combined index selection (CISr) in small family size (20) (Table 2, 1998b); unrestricted phenotypic selection was found to be similar to or better than restricted combined index selection at the same status

Table 1. – Gain and breeding population size of phenotypic selection (PS) and combined index selection (CIS) at the same status number with the same number of test plants.

Initial heritability	Generation	Status number (Ns)	Gain (G)		Breeding population size (N)		
			PS	CIS	PS	CIS	
0.01	1	5	2.49	5.62	9	37	
		10	2.41	4.73	19	58	
		20	2.10	3.68	40	90	
		40	1.81	2.84	79	135	
		80	1.55	2.16	155	209	
		100	1.50	2.00	191	243	
	5	200	1.30	1.45	369	400	
		5	10.89	17.44	29	91	
		10	10.33	14.31	59	131	
		20	8.91	11.22	116	193	
		40	7.30	8.64	222	295	
		80	5.75	6.43	409	446	
	0.20	1	100	5.07	5.54	494	517
			5	10.85	14.38	11	35
			10	10.44	13.22	23	59
			20	9.58	11.65	45	92
			40	8.44	10.03	88	139
			80	7.20	8.37	168	213
5		100	6.85	7.77	205	247	
		200	5.34	5.70	382	403	
		5	43.31	46.26	42	102	
		10	38.87	40.40	81	158	
		20	33.87	36.35	152	215	
		40	28.73	28.91	265	325	
0.50		1	80	22.55	23.25	447	467
			100	20.04	20.42	531	538
			5	16.53	17.84	13	28
			10	16.12	16.28	29	52
			20	14.54	14.91	56	86
			40	13.06	13.42	103	134
	5	80	11.20	11.54	187	211	
		100	10.57	10.81	225	246	
		200	8.17	8.24	400	404	
		5	61.47	57.61	52	97	
		10	55.34	52.09	99	151	
		20	47.94	47.23	174	215	
	1.0	1	40	40.13	39.21	295	329
			80	31.97	32.12	472	475
			100	28.33	28.57	557	547
			5	23.57	23.23	21	21
			10	21.94	21.76	43	42
			20	19.53	19.67	77	73
5		40	17.45	17.69	130	120	
		80	15.04	15.25	217	199	
		100	14.14	14.43	257	235	
		200	11.13	11.23	429	395	
		5	84.70	70.30	49	85	
		10	74.95	65.60	96	133	
5		20	64.79	59.84	172	201	
		40	53.96	51.29	295	317	
		80	42.84	42.59	476	468	
		100	37.88	38.30	563	543	

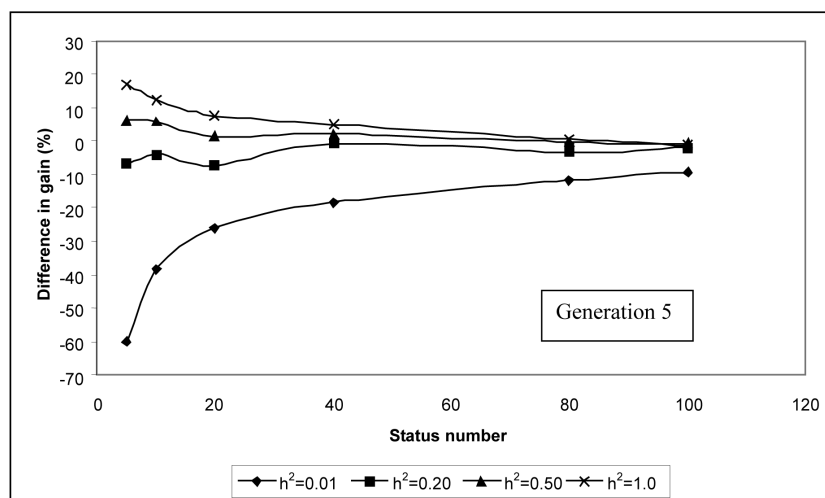
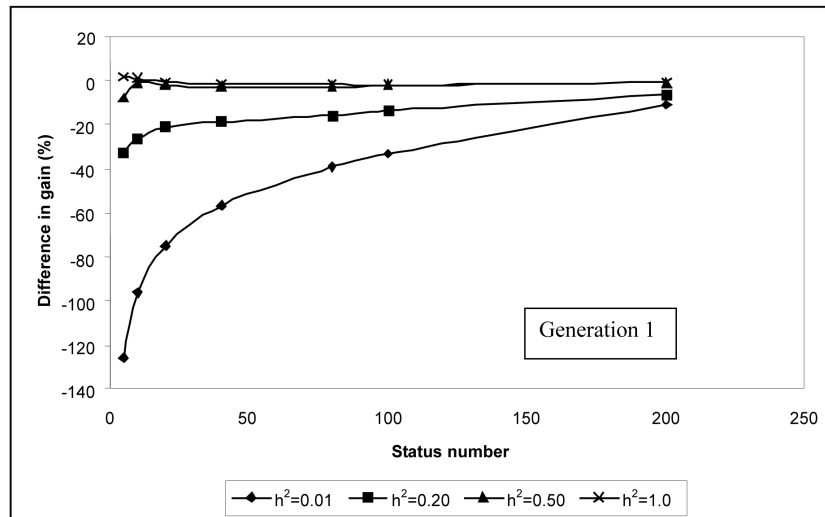


Figure 1. – Difference in gain (δG) between phenotypic and combined index selection at different heritability at the same status number. Positive difference means that phenotypic selection offers highest gain.

number for heritability higher than 0.20 (1998a). In principle CIS in the context used here is more efficient than CISr, so the later should be used with care or not at all. As an alternative where CISr is used, it should be considered if the size of the breeding population could be expanded, so CIS could be used leaving an acceptable level of gene diversity.

Phenotypic selection and combined index selection

Phenotypic selection has been shown to be a surprisingly efficient method for achieving gain while preserving genetic variance and gene diversity from generation to generation and is often compatible with restricted index selection (e.g. ANDERSSON et al., 1998), LINDGREN (2003) recommended phenotypic selection as an alternative low-input technique for seed production and long-term breeding. We found that, under constant plant number and at low heritability (0.01), the selection efficiency of phenotypic selection was much inferior to that of combined index selection, in particular in combination with low status number. Only when heritability was

high, phenotypic selection was as efficient as or, after multiple generations, slightly more efficient than combined index selection. In contrast, ANDERSSON (1998b) pointed out that, under constant breeding population size, at the same status number and in small family sizes, phenotypic selection could get the same gain (at low heritability) or more (at high heritability) compared with restricted combined index selection. In this study consideration was made only for family performance and individual performance. Real breeders nowadays use the information more efficient, particular in advanced generations, and if such breeding value estimates are used phenotypic selection will be less efficient (cf. the development of POPSIM used by LSTIBUREK, 2005).

Should tree breeders use combined index selection or phenotypic selection?

The resources were considered constant as the number of experimental plants was kept constant, but this is generally unfair against phenotypic selection. Phenotypic selection requires simple calculations; the risk for

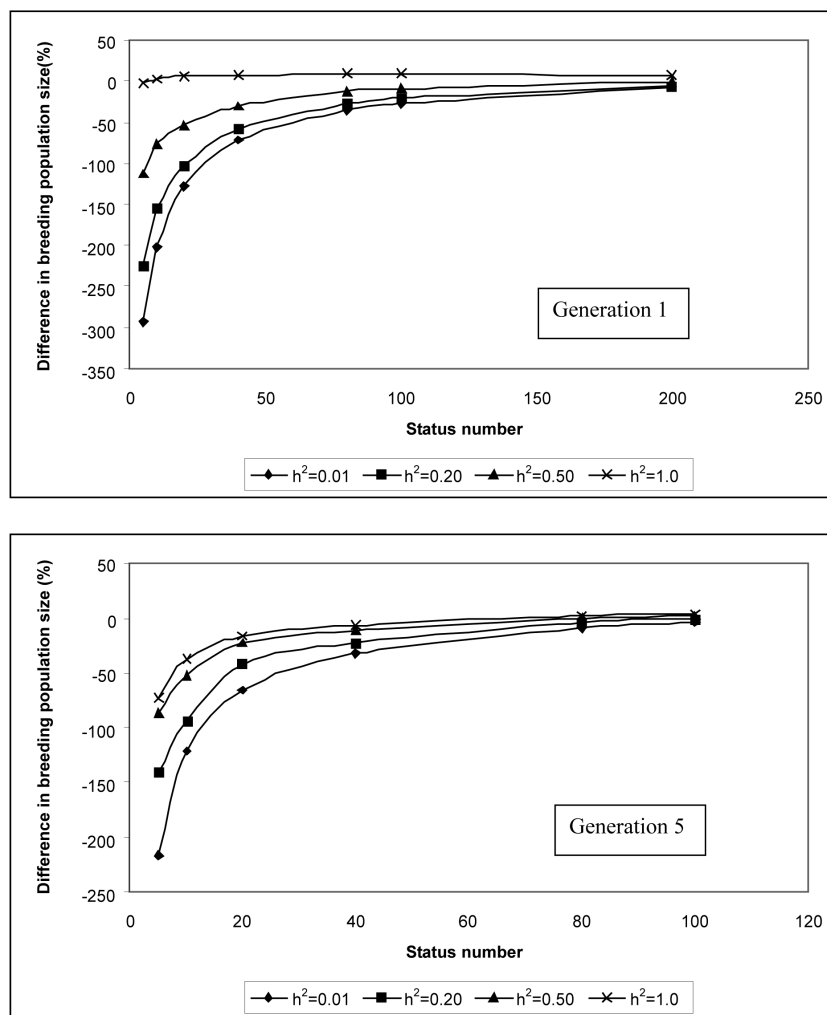


Figure 2. – Difference in population size (δN) (%) between phenotypic and combined index selection at different heritability at the same status number.

Table 2. – Comparison between unrestricted combined index selection (CIS) and restricted combined index selection (CISr) at initial heritability 0.20 (δG means percent superiority in gain).

generation	Status number	δG %		
		(PS-CIS)/PS	(PS-CISr)/PS	(CIS-CISr)/CIS
1	89	-16.31	7.87	20.79
5	22	-7.31	-0.77	6.11

The following data were obtained from Andersson 1998b (Table 3), family size 20.

1	89	0.0
5	22	-3.0

mistake is small; not even field identification of experimental plants are needed; the selection can be made faster as no complicated in office calculations are needed. It would be fair to put a rather high price on the size of the breeding population. Pollen management must be done on more individuals, more artificial crosses need to be made, clone archives become larger. Typically, the breeding population size is not controlled by the person choosing the selection strategy, and the breeding population may often have been set too low to use CIS without unacceptable gene diversity losses. If large breeding populations cannot be conveniently used, that is an argument for using phenotypic selection. For reasonable large heritability the difference between the selection

alternatives in our model is almost negligible, in particular after some generations and if low status number is not tolerated. Thus other factors outside the model may be decisive and make phenotypic selection more favourable. Thus, it is often justified to consider phenotypic selection as an alternative to combined index selection.

Implication for the size of the breeding population.

The optimal size of the breeding population is generally high (Table 1), in particular considering the low resource in this study, and in particular when associated with combined index selection. This is an argument for using rather large breeding populations.

Acknowledgements

The work was done while LI HUOGEN was a visiting scholar at Swedish University of Agricultural Science in Umeå. We gratefully thank Dr. TIM MULLIN and MILAN LSTIBŮREK for support with the breeding simulator POP-SIM including a version that accommodates this study and their comments. We also gratefully acknowledge financial support from the Kempe Foundation and the China Scholarship Council (CSC).

Literature cited

- ANDERSSON, E. W., K. A. SPANOS, T. J. MULLIN and D. LINDGREN (1998a): Phenotypic selection can be better than selection for breeding value. *Scand. J. For. Res.* **13**: 7–11.
- ANDERSSON, E. W., K. A. SPANOS, T. J. MULLIN and D. LINDGREN (1998b): Phenotypic selection compared to restricted combined index selection for many generations. *Silva Fennica* **32**(2): 111–120.
- BURTON, G. W. (1974a): Recurrent restricted phenotypic selection increases forage yields of *Pensacola bahiagrass*. *Crop Sci.* **14**: 831–835.
- BURTON, G. W. (1974b): Improved recurrent restricted phenotypic selection increases *Bahiagrass* forage yields. *Crop Sci.* **22**: 1058–1061.
- COCKERHAM, C. C. (1967): Group inbreeding and coancestry. *Genetics* **56**: 89–104.
- COTTERILL, P. P. (1986): Genetic gains from alternative breeding strategies including simple low cost options. *Silvae Genet.* **35**: 212–223.
- DANUSEVIČIUS, D. and D. LINDGREN (2005): Optimisation of breeding population size for long-term breeding. *Scandinavian Journal Forest Research* (20) **1**: 18–25.
- DEMPFLE, L. (1990): Problems in the use of the relationship matrix in animal breeding. *In: Advances in statistical methods for genetic improvement of livestock.* (ed. D. GIANOLA and K. HAMMOND). Pp. 454–473. Springer-Verlag, Berlin.
- FALCONER, D. S. (1989): *Introduction to Quantitative Genetics*, 3rd edition. Longman, Scientific and Technical.
- LINDGREN, D., L. D. GEA and P. A. JEFFERSON (1996): Loss of genetic diversity monitored by status number. *Silvae Genetica* **45**: 52–59.
- LINDGREN, D. (2003): Low-input tree breeding strategies. *In: Eucalyptus plantations-Research, Management and Development*, R.-P. WEI and D. XU (eds), World scientific, Singapore, 146–149.
- LSTIBŮREK M, 2005: Population response to positive assortative mating in forest tree breeding, PhD dissertation. North Carolina State University, Raleigh. USA.
- MULLIN, T. J. and Y. S. PARK (1995): Stochastic simulation of population management strategies for tree breeding: a new decision support tool for tree personal computers. *Silvae Genet.* **44**: 132–141.
- ROSVALL, O. (1999): Enhancing gain from long-term forest tree breeding while conserving genetic diversity. Ph.D Thesis. Swedish University of Agricultural Sciences, Department of Forest Genetics and Plant Physiology, Umeå.
- WEI, R.-P. and D. LINDGREN (1991): Selection effects on diversity and gain. *Silva Fennica* **25**(4): 229–234.
- WEI, R.-P. and D. LINDGREN (1993): Phenotypic selection was more efficient than combined index selection when applied on full sibs of lodgepole and Scots pine. *Dept. of Forest Genetics and Plant Physiology, SUAS Report* **11**: 289–292.

Phenotypic Correlation and Path Coefficient Analysis of Nut Yield and Yield Components in Cashew (*Anacardium occidentale* L.).

By O. M. ALIYU

Department of Plant Breeding, Cocoa Research Institute of Nigeria, P M B 5244, Ibadan, Nigeria
walealiyu@mail.com

(Received 4th July 2005)

Abstract

Relationships between cashew nut yield and nine agronomic traits comprising seven reproductive (nut and floral) and two vegetative characters were studied in 59 selected cashew genotypes over three production seasons. Phenotypic correlation analysis showed that nuts per panicle ($r = 0.844$), number of nuts per tree ($r = 0.988$) and number of hermaphrodite flowers per panicle ($r = 0.863$) were positively correlated with nut yield and could be used as primary components for improving yield. Although correlation analysis showed insignifi-

cant association between nut weight and nut yield, path analysis revealed that the trait had significant positive direct effect (0.317) on nut yield. The subtle indirect effects of nut weight and leaf size on nut yield were more important than their direct effects and could be classified as secondary components. Both the direct and indirect effects of weight of the whole fruit and tree canopy on nut yield were negative and appeared detrimental.

Key words: *Anacardium occidentale*, Correlation analysis, Path analysis, Breeding.